**一、上次课堂作业分析**

* **9道判断题，还是有很多同学做错，请大家仔细学习解析，掌握自己为什么做错。通过这几道判断题说明同学们对于关系代数的知识还是掌握不够的，需要加强复习。最终对大家的要求是给你一个查询任务的题目的叙述，你就要直接把正确的关系代数的查询式子写出来。**
* **那什么是正确的关系代数的查询式子呢？通过这9道题目大家要明白，关系和关系模式是不一样的概念，关系模式是 ‘型’的概念，是静态的；而关系是‘值’的概念，是动态的，关系表的具体值都是可以变化的，但是无论怎么变它都需要满足一个关系模式的约束，关系模式可以体现完整性的约束条件（主码是什么？外码是什么？等等），正确的关系代数的式子是在某种关系模式下，无论你的关系表如何变，你都能通过这个关系代数的式子将正确的结果查询出来。**
* **关系代数答案不唯一，关系代数中属性、表名要按照原本题目中的样子书写，即区分大小写，注意属性的分量值是否属于字符类型，若属于字符类型，判断时应当打单引号。**

1. 课件例2.10中查询至少选修1号课程和3号课程的学生号码。

该关系代数表达式正确么？

（∏Sno(σCno=‘1’(SC))）∩（∏Sno(σCno=‘3’(SC))）

对，∏Sno(σCno=‘1’(SC))得到的是选了1号课程的学生的学号，∏Sno(σCno=‘3’(SC))得到的是选了3号课程的学生的学号，两者求交就为至少选修了1号和3号课程的学生号码，为什么满足 “至少”的条件呢？例如，如果是选了1、3、4号课程的学生学号也会被该关系代数的式子查询到，是符合题目要求的。

2. 课件例2.10中查询至少选修1号课程和3号课程的学生号码。

该关系代数表达式正确么？

∏Sno（σCno=‘1’(SC)∩σCno=‘3’(SC)）

错，比如

σCno=‘1’(SC) 查到 表1

|  |  |  |
| --- | --- | --- |
| Sno | Cno | Grade |
| 2001 | 1 | 90 |

σCno=‘3’(SC) 查到 表2

|  |  |  |
| --- | --- | --- |
| Sno | Cno | Grade |
| 2001 | 3 | 80 |

根据表1与表2，σCno=‘1’(SC)∩σCno=‘3’(SC)为空集，请同学们回头看R∩S的计算原则。但是根据题意这个2001号同学应该是被查询出来的。

通过对比题目1与题目2，这就是为什么题目一在求交之前要做投影处理：即对原表符合条件（选了某课程）的元组筛选出来之后再对Sno进行投影。

3. 课件例2.10中查询至少选修1号课程和3号课程的学生号码。

该关系代数表达式正确么？

∏Sno，Cno(SC) ÷ ∏ Cno(σCno=‘1’**∨**Cno**=**‘3’ (Course))

对，课件与教材例2.10的答案就为∏Sno，Cno(SC) ÷K，而临时表K为

|  |
| --- |
| Cno |
| 1 |
| 3 |

∏ Cno(σCno=‘1’**∨**Cno**=**‘3’ (Course))的含义等价于建立临时表K，针对Course表选择Cno值满足1号或3号的元组，这些元组是包含Course表中所有的属性的，所以再进行Cno的投影。投影是要去掉重复的，所以最终得到的结果就是等价于临时表K。

4. 课件例2.10中查询至少选修1号课程和3号课程的学生号码。

该关系代数表达式正确么？

∏Sno，Cno(SC) ÷ ∏ Cno(σCno=‘1’**∧**Cno**=**‘3’ (Course))

错，因为σCno=‘1’**∧**Cno**=**‘3’ (Course)表示：对于Course表中选择逻辑表达式“Cno=‘1’**∧**Cno**=**‘3’”为真的元组，但是对于每一个元组的Cno的值都只可能有一个值，不可能既等于1又等于3。因此该关系代数式错误。

5. 课件例2.10中查询至少选修1号课程和3号课程的学生号码。

该关系代数表达式正确么？

∏Sno(σCno=‘1’**∧**Cno**=**‘3’(SC))

错，与第4题错误原因一致，那同学们会问如果把**∧改为∨呢？**

∏Sno(σCno=‘1’**∨**Cno**=**‘3’(SC))也为错，因为该表达式不能保证选出来的Sno是至少选择了1号和3号课程的，只能保证查询到选了1号课程或者3号课程的学生的学号。这也是课件例2.10为什么要用÷，因为除法表示包含的意思，可以体现出“至少”。

6. 课件例2.10中查询至少选修1号课程和3号课程的学生号码。

该关系代数表达式正确么？

∏Sno (SC ÷ ∏ Cno(σCno=‘1’**∨**Cno**=**‘3’ (SC)))

错，∏ Cno(σCno=‘1’**∨**Cno**=**‘3’ (SC))与课件及教材的例2.10中的临时表K是等价的。那么我们看看SC÷K等于什么？

假设SC表只有2001学号同学，如表3，2001学号的学生选了1、3、4号课程，2001号同学肯定是查询结果，但是我们看如果是SC÷K=Φ（空集）。SC表中（2001,80）的象集为1，（2001,75）的象集为3，（2001,90）的象集为4。没有任何一个(Sno,Grade)值的象集包含K表对共有属性Cno的投影结果（1,3）。所以，SC÷K=Φ，为空集。该题的核心考察大家对除法运算的掌握。

表3

|  |  |  |
| --- | --- | --- |
| Sno | Cno | Grade |
| 2001 | 1 | 80 |
| 2001 | 3 | 75 |
| 2001 | 4 | 90 |

7. 查询选修了全部直接先行课为6号课程的学生姓名。

该关系代数表达式正确么？

∏ Sname(σCpno=‘6’(Course）⋈SC⋈∏ Sno,Sname (Student))

错，题目的核心为全部直接先行课为6号课程，跟课件和教材的例2.12是不一样的。题目的意思是：假设以课程6为先行课的所有课程为课程3、课程4。所求的学生应当是至少同时选修了课程3与课程4的。而题目中的式子不能完成这个要求，题目中只是表达查询至少选修了一门先行课为6号课程的同学。该题正确的关系代数表达式为（课件上也有）

∏ Sname((∏ Sno,Cno (SC) ÷ ∏ Cno (σCpno=‘6'(Course))) ⋈ Student)

8. 查询选修了全部直接先行课为6号课程的学生姓名。

该关系代数表达式正确么？

∏ Sname(σCpno=‘6’(Course）⋈∏ Sno,Sname (Student) ⋈SC)

错，除了第7题的错误点外，σCpno=‘6’(Course）⋈∏ Sno,Sname (Student)是不能进行运算的，因为没有共同的属性，而自然选择需要具有同名属性。注意σCpno=‘6’(Course）的属性组与原Course一致。

9. 查询选修了全部直接先行课为6号课程的学生姓名。

该关系代数表达式正确么？

(∏ Sname,Cno (Student⋈SC)) ÷(∏ Cno (σCpno=‘6'(Course)))

错，假设以课程6为先行课的所有课程为课程3、课程4，∏ Cno (σCpno=‘6'(Course))等价于建立临时表S，如表4。

表4

|  |
| --- |
| Cpno |
| 3 |
| 4 |

假设Student为表5：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sno | Sname | Ssex | Sage | Sdept |
| 2001 | 李勇 | 男 | 17 | MA |
| 2004 | 李勇 | 男 | 18 | MA |

假设SC为表6：

|  |  |  |
| --- | --- | --- |
| Sno | Cno | Grade |
| 2001 | 3 | 80 |
| 2004 | 4 | 75 |

Student⋈SC为表7：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sno | Sname | Ssex | Sage | Sdept | Cno | Grade |
| 2001 | 李勇 | 男 | 17 | MA | 3 | 80 |
| 2004 | 李勇 | 男 | 18 | MA | 4 | 75 |

∏ Sname,Cno (Student⋈SC)为表8：

|  |  |
| --- | --- |
| Sname | Cno |
| 李勇 | 3 |
| 李勇 | 4 |

∏ Sname,Cno (Student⋈SC) ÷S为

|  |
| --- |
| Sname |
| 李勇 |

但是大家可以看到叫李勇的有2位同学，实际上任何一位李勇同学都没有同时修课程3、4，按理说不符合题目的要求的，但是最后的查询结果却将其查出来了。所以本题表达式错误，这也就是为什么正确答案（见习题7）中在除法的左端为∏ Sno,Cno (SC)，这里的投影属性含有主属性Sno。Sname并不是Student表的主码。所以为什么先讲完整性约束的概念再来到关系代数的讲解，所有知识点一环套一环。整个书本知识点是由浅入深。

**二. 本次课的课堂任务**

本次课继续学习SQL语言，本次课将开始学习SQL最重要的知识点：数据的查询。

1. 今天首先学习视频P18-P20，网址为

<https://www.bilibili.com/video/BV13J411J7Vu?p=18>

涉及教材的P89-103面，配套学习课件“数据库系统概论-计科-第三章\_2.pdf”

今天涉及两个部分：单表查询、连接查询。

* 首先从最简单的单一关系表的查询开始学起SQL查询的语法，分为：

1）选择表中的若干列：\*代表什么含义？对属性名怎么采用别名的方法？

2）选择表中的若干元组： 用什么关键词能达到关系代数中∏的作用？比较大小运算符，IN\NOT IN，字符匹配查询中LIKE、ESCAPE、%、\_ 都分别是什么含义？怎么使用？空值查询和符合逻辑表达式查询怎么书写？

3）ORDER BY子句：什么时候用？默认为正序还是倒序？只能放在查询语句的什么位置？

4）聚集函数：聚集函数有哪些？都有什么含义？什么时候使用？

5）GROUP BY子句：课件中扩充了一个书上及视频中没有的例题，请注意学习。

SELECT相当于对属性的操作，即关系表中的列提取。WHERE相当于对满足条件的行操作，类似于关系代数中的σ。聚集函数与GROUP BY子句是难点，什么是聚集函数？GROUP BY 经常与HAVING 连用，大家要重点区分WHERE和HAVING 的作用对象，都是满足条件的筛选出来，但是两者的作用对象有什么区别？WHERE子句中能出现聚集函数么？

* 如果查询涉及到2个及2个以上的表时，就可能面临着连接查询。

任意两个表进行连接查询的执行步骤是什么？有哪两种类型？怎样用SQL语言表达关系代数中R⋈S？有些查询任务必须通过自身表的连接进行，那对自身表进行连接时，是将其当做不同的表还是相同的表进行操作呢？自身表连接与不同的两个表连接在执行步骤原理上是一样的么？3个及3个以上的表连接的关键WHERE的筛选条件一般需要怎么写？